

The Analysis of China CPI Trend Forecast Based on ARIMA Model

Meizhen Liu^{1, a, *}, Chunmei Duan^{1, b}

¹Shandong Normal University, Jinan250000, China.

^a18353361502@163.com, ^bcmduan@sdnu.edu.cn

Abstract. The Consumer Price Index (CPI) is an index of price changes that is calculated based on the prices of products and services related to residents' lives. This paper selects China's latest monthly CPI data as the research object, and constructs an ARIMA (12, 1, and 20) model of autoregressive-rendered moving average. Based on the excellent model fitting effect, it successfully predicts future CPI trends. The effective implementation of price control policies at the macro level provides a quantitative basis.

Keywords: CPI, ARIMA, forecast, auto regression.

1. Introduction

In recent years, the rising prices of agricultural products have become a topic of common concern among the people and economists. The Consumer Price Index (CPI) is a comprehensive indicator of the prices of products and services related to the lives of residents. It can be used to measure changes in prices. The price index showed an upward trend, indicating that macroeconomic inflation has constrained the consumer price index of residents to be constrained by a variety of factors, and the factors have maintained an extremely complicated relationship. Especially in the post-international financial crisis era, China's economy is not operating at all. Certainty and instability are on the rise. Faced with this situation, the market's voice of concern about the subsequent development of the Chinese economy continues to permeate. In this context, the "uncertainty" of the impact factor has increased. Therefore, using the structural causal model to predict the trend of consumer price index, it is generally difficult to achieve a more satisfactory forecast effect. The ARIMA model is one of the most widely used linear time series modeling tools. The ARIMA model is also called the Box-Jenkins Model (B-J model). It is a time series prediction method named after the US statistician George E.P. Box and the British statistician GwilymM. Jenkins. This paper proposes to use the latest monthly CPI data from January 2001 to October 2011 as research objects to construct the ARIMA (12, 1, and 20) model, and predict the future trend of CPI on the basis of good model fitting effect. In order to provide a quantitative basis for government decision-making departments to effectively implement price control policies.

2. The Model Setting

Time series forecasting uses historical data to analyze the internal rules of the change of target objects over time, and then uses extrapolation mechanisms to extrapolate this rule to the future; that is, to process the time series to study the changing trends of the target itself. In order to accurately predict the future changes in the target object. The basic idea of the analysis of time series is that some time series can be regarded as variables that change randomly with time t . Although the individual sequence values of the time series are uncertain, the whole sequence shows a certain variation pattern. The mathematical model is described approximately. In the real world, people often use the time series ARIMA model to conduct empirical research to achieve the optimal prediction effect in the sense of minimum variance. The ARIMA model is called the summation autoregressive moving average model, abbreviated as ARMA (p , d , and q). The model structure is as follows:

$$\begin{cases} \Phi(B)\nabla^d x_t = \Theta(B)\epsilon_t \\ E(\epsilon_t) = 0 \quad \text{var}(\epsilon_1) = \sigma_s^2, E(\epsilon_t, \epsilon_s) = 0, s \neq t \\ E x_s \epsilon_t = 0 \quad \forall s < 185 \end{cases}$$

3. Data Preprocessing

3.1 Chart Analysis

This article uses the latest monthly CPI from January 2008 to October 2017 in China (previous year = 100). The data (see Figure 1 for details) is from the official website of the National Bureau of Statistics <http://www.stats.gov.cn/>, All calculations are done using EViews 7.2 software.

Table 1. China's Consumer Price Index from 1978 to 2011 (1983 = 100)

years	CPI	years	CPI	years	CPI
1983	100	1996	216.4	2008	433.5
1984	101.9	1997	223.8	2009	438.7
1985	109.5	1998	238.1	2010	455.8
1986	112.2	1999	273.1	2011	464
1987	114.4	2000	339	2012	471
1989	116.7	2001	396.9	2013	493.6
1990	119.9	2002	429.9	2014	522.7
1991	131.1	2003	441.9	2015	519
1992	139.6	2004	438.4	2016	536.1
1993	149.8	2005	432.2	2017	565
1994	177.9	2006	434		
1995	209.9	2007	437		

From the perspective of changes in the annual CPI time series, the price index has a clear tendency to fluctuate over time. The initial judgment is a non-stationary time series. To further determine the stationarity of the time series, the unit root test was performed on the data using the ADF test method. The ADF test results show that the t-test value of CPI is greater than the street value, and this sequence can be considered as a significant non-stationary sequence. The CPI takes the first-order difference and then performs trend conversion. From the time trend of the CPI first-order differential sequence, the sequence is a purely random sequence. The ADF unit root test was performed on the D(CPI) sequence. The results showed that the t value was less than the critical value and the associated probability was significant. This sequence was a non-white noise sequence. It can be seen that the data after differential processing has been called a smooth non-white noise sequence, and the model can be fitted.

3.2 Diagram Analysis

Using sentences to generate a sequence of time points, draw a relationship diagram with time, that is, a sequence diagram, we can intuitively see the trend of increasing time, the average value of different time periods is different, there is a continuous rise, that is, the sequence is not stable.

Of course, this intuitive illustration is often misleading, so further judgment is needed.

Sample autocorrelation test. The sample correlation function graph available, as shown in Figure 1.

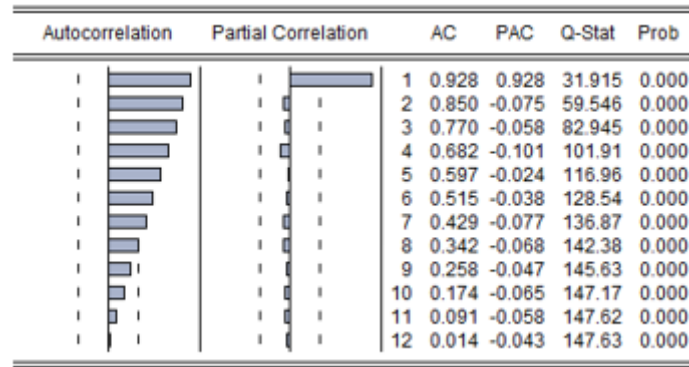


Figure 1. Sample autocorrelation test

The sample autocorrelation function for a time series is defined as:

$$r_k = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2}, \quad k = 1, 2, 3, \dots$$

It is easy to know that with the increase, the sample autocorrelation function decreases and tends to zero. However, in terms of the rate of decline, the stationary sequence is much faster than the non-stationary sequence. From the above sample correlation function graph, we can see that the sample correlation function is slowly decreasing towards zero, and it does not quickly decrease to zero like the partial autocorrelation function. Therefore, through the sample correlation diagram, it can be initially determined that the time series is non-stationary.

Of course, this method of judgment is also subjective. We will judge the objectivity in the following and further clarify the stationarity of the sequence.

A unit root test that tests the stability of the sequence is used. Get the result as shown in the figure2:

		t-Statistic	Prob.*	
Augmented Dickey-Fuller test statistic		-2.979579	0.1532	
Test critical values:	1% level	-4.273277		
	5% level	-3.557759		
	10% level	-3.212361		
Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPI(-1)	-0.166860	0.056001	-2.979579	0.0059
D(CPI(-1))	0.722663	0.127917	5.649460	0.0000

Figure 2. Unit root test

It can be seen that the probability $P = 0.1532$, and at the 5% level, the original hypothesis with the unit root is accepted.

The estimated result of Model 1 is:

$$\Delta CPI_t = -0.166860CPI_{t-1} + 0.722663\Delta CPI_{t-1} + 2.656493$$

In summary, according to the test, there is a unit root in the sequence, that is, the sequence is non-stationary.

4. Estimation Method

The known sequence CPI is a single-order integer, so we can estimate the model of the first-order differential sequence $ARMA(p, q)$, that is $ARIMA(p, 1, q)$, the model of the original sequence. The order $ARMA(p, q)$ is determined by looking at the sample's autocorrelation function and partial autocorrelation function. Therefore, we first need to draw the autocorrelation function and partial autocorrelation function of the sample CPI of the first-order differential sequence.

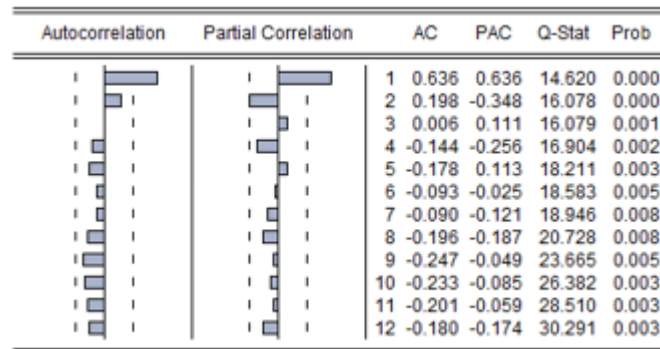


Figure 3. The sample CPI of the first-order differential sequence

We can see that the 1st order and 2nd order of the partial autocorrelation function graph have obvious sharp pillars, while the 1st order of the autocorrelation function graph has obvious sharp pillars, so we might as well preliminarily set the order as $p = 1$ $q = 2$. $ARIMA(2, 1, 1)$ The original sequence is identified.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.26277	39.93803	-0.632549	0.5328
D(CPI(-1))	2.686492	4.222884	0.636175	0.5304
D(CPI(-2))	-0.954266	1.758106	-0.542781	0.5921
MA(1)	0.074057	0.737033	0.100480	0.9208
RESID(-1)	-2.767403	4.417799	-0.626421	0.5367
RESID(-2)	0.107175	1.133900	0.094519	0.9255
R-squared	0.016325	Mean dependent var		0.035445
Adjusted R-squared	-0.180410	S.D. dependent var		11.76864
S.E. of regression	12.78624	Akaike info criterion		8.106602

Figure 4. Original sequence of CPI

After obtaining the estimate, we also perform a 1-order LM test on the sequence residuals to obtain the 1-order LM test results of the sequence residuals as shown in Fig. 5.

The estimated results of the model according to Figure 5 are:

$$\Delta CPI_t = -25.26277 + 2.686492\Delta CPI_{t-1} - 0.954266\Delta CPI_{t-2} + \varepsilon_t + 0.074057\varepsilon_{t-1}$$

$$AIC=8.106602 \quad SC=8.384147 \quad LM(1) = 0.361366$$

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-25.26277	39.93803	-0.632549	0.5328
D(CPI(-1))	2.686492	4.222884	0.636175	0.5304
D(CPI(-2))	-0.954266	1.758106	-0.542781	0.5921
MA(1)	0.074057	0.737033	0.100480	0.9208
RESID(-1)	-2.767403	4.417799	-0.626421	0.5367
RESID(-2)	0.107175	1.133900	0.094519	0.9255
R-squared	0.016325	Mean dependent var		0.035445
Adjusted R-squared	-0.180410	S.D. dependent var		11.76864
S.E. of regression	12.78624	Akaike info criterion		8.106602

Figure 5. The 1-order LM test results of the sequence residuals

It can be seen that the residuals of the two models are irrelevant, that is, the residuals can be considered as random sequences, that is, the models are basically consistent. In summary, we finally select the model $ARIMA(1,1,1)$ to estimate the sequence CPI , the caliber of the model that gets the sequence CPI is:

$$\Delta CPI_t = -25.26277 - 2.626492CPI_{t-1} - 0.954266CPI_{t-2} + \varepsilon_t + 0.074057\varepsilon_{t-1}$$

$$Var(\hat{\sigma}_\varepsilon^2) = 19.048598$$

5. Forecast and Analysis

First, in order to test the effect of model prediction, static in-sample prediction was performed using Views 6.0 software, and then the predicted values and real values were plotted as corresponding graphs. The real value of China's monthly CPI during the sample period from Figure 4 (blue in the figure The color line represents the CPI) and the predicted value (red lines in the figure Table, namely CPIF) The fitting of the two curves, the trend of the two is basically consistent, and the similarity in many places indicates that the model has good fitting accuracy and the corresponding prediction accuracy is also high. According to the comparison between the predicted values and the real values for January-June 2017 as listed in Table 2, the prediction errors are all in the range of 0.5 to 10,000.

The prediction accuracy is very high.

To further analyze the trend of China's CPI in 2018, reuse the established table.2.

Table 2. Intra-sample predictions for China's monthly CPI

2018 years	1	2	3	4	5	6
actual value	104.9	104.9	105.0	105.1	105.2	105.4
Predictive value	105.35	04.59	105.10	104.93	04.8	105.2
Prediction error (‰)	-0.43	0.30	-0.0	0.16	0.29	0.11

The ARIMA (12, 1, and 20) model performs out-of-sample dynamic forecasting of the CPI index and extends the modeled sample from January 2006 to December 2017. The medium-term forecast of the data can be obtained directly by running the Views software. Also plot the CPI real and predicted values as a polyline, from the CPI plotted in Figure 5.

From the point of view of true and dynamic forecast values, it can be seen that in the period from now to December 2017, China's price trend still has great difficulty in regulation, and relevant departments need to take effective measures to strictly control price increases.

Acknowledgements

This research was supported by the National Nature Science Foundation of China (61502284, 61602282).

References

- [1]. Yao Ting. Analysis and Prediction of the Macro-factors of Stock Prices [D]. Bengbu: Anhui University of Finance and Economics, 2014 (05).
- [2]. You Zuojun. Research and Application of Time Series Analysis in Stock [D]. Shenyang: Shenyang University of Technology, 2014 (02).
- [3]. Li Yoking, Cheng Longman. The Application of Time Series Model on Stock Price Forecast [J]. Market Modernization, 2011 (11).

- [4]. Zhang Chao. Stock Price Forecast Based on ARMA-GARCH Model [J]. Journal of Nanjing University of Aeronautics and Astronautics, 2014 (09).
- [5]. Dong Bolin, Xu Dungy. Prediction and Analysis of Stock Price of Agricultural Products Based on ARIMA Model [J]. Modern Business, 2015 (03).
- [6]. Gao Yuan. Empirical Study on LETV Stock Price Forecast Based on ARMA Model [J]. Modern Economic Information, 2015 (07).
- [7]. Sun Xinxiang. The Application of Time Series Model in Stock Price Forecasting [D]. Kunming: Yunnan University, 2016 (05).
- [8]. Zhang Nan. Research on Stock Trend Forecasting Based on Time Series and R Language Application [J]. Modern Business, 2016 (08).
- [9]. Wu Wuxia, Wen Xin. Short-term Stock Price Forecasting Based on ARIMA Model [J]. Statistics & Decision, 2016 (12).
- [10]. Ma Yana, Zeng Joying. Prediction and Analysis of the Shanghai Composite Index Based on the ARIMA Model [J]. Economic and Trade Practice, 2017 (02). Journal of Hunan University of Arts and Science (Natural Science Edition), 2017 (09).
- [11]. Zhang Jibe. Analysis of Volatility of Stock Price Based on Time Series Model [J]. Journal of Hunan University of Arts and Science, 2017 (09).